BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 5: e75506
doi: 10.3897/biss.5.75506

OPEN ACCESS

Conference Abstract

The Verification of Ecological Citizen Science Data: Current approaches and future possibilities

Emily Baker[‡], Jonathan P Drury[‡], Johanna Judge[§], David B. Roy[|], Graham C Smith[¶], Philip A Stephens[‡]

‡ Conservation Ecology Group, Department of Biosciences, Durham University, Durham, United Kingdom
§ National Biodiversity Network Trust, Nottingham, United Kingdom
| Centre for Ecology & Hydrology, Crowmarsh Gifford, Wallingford, Wallingford, United Kingdom
¶ National Wildlife Management Centre, Animal and Plant Health Agency, York, United Kingdom

Corresponding author: Emily Baker (emily.baker@durham.ac.uk)

Received: 20 Sep 2021 | Published: 20 Sep 2021

Citation: Baker E, Drury JP, Judge J, Roy DB, Smith GC, Stephens PA (2021) The Verification of Ecological Citizen Science Data: Current approaches and future possibilities. Biodiversity Information Science and Standards 5: e75506. https://doi.org/10.3897/biss.5.75506

Abstract

Citizen science schemes (projects) enable ecological data collection over very large spatial and temporal scales, producing datasets of high value for both pure and applied research. However, the accuracy of citizen science data is often questioned, owing to issues surrounding data quality and verification, the process by which records are checked after submission for correctness. Verification is a critical process for ensuring data quality and for increasing trust in such datasets, but verification approaches vary considerably among schemes. Here, we systematically review approaches to verification across ecological citizen science schemes, which feature in published research, aiming to identify the options available for verification, and to examine factors that influence the approaches used (Baker et al. 2021). We reviewed 259 schemes and were able to locate verification information for 142 of those. Expert verification was most widely used, especially among longer-running schemes. Community consensus was the second most common verification approach, used by schemes such as Snapshot Serengeti (Swanson et al. 2016) and MammalWeb (Hsing et al. 2018). It was more common among schemes with a larger number of participants and where photos or video had to be submitted with each record. Automated verification was not widely used among the schemes reviewed. Schemes that used automation, such as eBird (Kelling et al. 2011) and Project FeederWatch (Bonter and

© Baker E et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Cooper 2012) did so in conjunction with other methods such as expert verification. Expert verification has been the default approach for schemes in the past, but as the volume of data collected through citizen science schemes grows and the potential of automated approaches develops, many schemes might be able to implement approaches that verify data more efficiently. We present an idealised system for data verification, identifying schemes where this hierachical system could be applied and the requirements for implementation. We propose a hierarchical approach in which the bulk of records are verified by automation or community consensus, and any flagged records can then undergo additional levels of verification by experts.

Keywords

ecology, big data, verification, crowdsourcing, data quality

Presenting author

Emily Baker

Presented at

TDWG 2021

Acknowledgements

EB was supported by the Natural Environment Research Council's IAPETUS2 Doctoral Training Partnership award number NE/S007431/1. DR was supported by the Natural Environment Research Council award number NE/R016429/1 as part of the UK-SCAPE programme delivering National Capability.

References

- Baker E, Drury J, Judge J, Roy D, Smith G, Stephens P (2021) The Verification of Ecological Citizen Science Data: Current Approaches and Future Possibilities. Citizen Science: Theory and Practice 6 (1). https://doi.org/10.5334/cstp.351
- Bonter DN, Cooper CB (2012) Data validation in citizen science: a case study from Project FeederWatch. Frontiers in Ecology and the Environment 10 (6): 305-307. https://doi.org/10.1890/110273
- Hsing P, Bradley S, Kent V, Hill R, Smith G, Whittingham M, Cokill J, Crawley D, Stephens P, (2018) Economical crowdsourcing for camera trap image classification. Remote Sensing in Ecology and Conservation 4 (4): 361-374. https://doi.org/10.1002/rse2.84

- Kelling S, Yu J, Gerbracht J, Wong W (2011) Emergent Filters: Automated Data Verification in a Large-Scale Citizen Science Project. 2011 IEEE Seventh International Conference on e-Science Workshops https://doi.org/10.1109/esciencew.2011.13
- Swanson A, Kosmala M, Lintott C, Packer C (2016) A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. Conservation Biology 30 (3): 520-531. https://doi.org/10.1111/cobi.12695